BLOMST – An Optimization Model for the Bioenergy Supply Chain

Michal Kaut Ruud Egging Truls Flatberg
Kristin Tolstad Uggen

June 6, 2014

In this chapter, we present a new model for optimal strategic and tactical planning of the bioenergy supply chain under uncertainty. We discuss specific challenges, characteristics and issues related to this type of model. The technological details, variability in supply and demand, and uncertainty in virtually all aspects of the supply chain require advanced modeling techniques. Our model provides a broad modeling approach that addresses the entire supply chain using an integrated perspective.
The broad applicability of the approach is illustrated by the two cases discussed at the end of the chapter. The first case presents a forest to bioenergy supply chain in a region of the Norwegian west coast. The second case presents the miscanthus supply chain to a transformation plant in Burgundy, France and takes into consideration uncertain final demand.

1 Introduction

Renewable energy is the fastest-growing source of energy generation, according to the IEO2013 Reference case (U.S. Energy Information Agency, 2013). Total power generation from renewables is projected to increase by 2.8% annually until 2040. Although about 80% of the total increase is in hydroelectric and wind power, bioenergy generation is expected to grow at about the same pace. The growth of non-hydro renewable energy sources in OECD Europe is stimulated by renewable energy policies. The EU mandates that 20% of total energy production must come from renewables by 2020 (according to the 20-20-20 target); this is up from about 13% in 2010. In addition to the EU targets, several countries provide incentives that promote the expansion of renewable energy generation. For example, Germany, Spain, Denmark, and the United Kingdom have enacted feed-in tariffs that guarantee minimum prices for energy generated from renewable sources.

The market for bioenergy plants is expected to grow quickly in the coming years (ecoprog GmbH, 2013). In 2020, 3500 bioenergy plants are expected to be operational worldwide, implying a growth of installed capacity by about 46% in an eight-year period.

A beneficial characteristic of biomass is that it can be stored, in contrast with solar or wind-based generation, which means that bioenergy generation can more easily be matched with varying demand. Unfortunately, biomass often requires large production and collection areas, has low energy density, is expensive to harvest and transport, and has high maintenance and logistics costs. This makes it challenging for the bioenergy industry to compete with the highly developed fossil-fuel value chains (De Meyer, Dirk, Rasinmäki, and Van Orshoven, 2014). Also, biomass quality can be affected by transport and storage. Although *passive* drying reduces the moisture content, which is a favorable outcome, storage can also induce fiber deterioration, which reduces the energy content (Wolfsmayr and Rauch, 2014). Furthermore, bioenergy supply chains have to deal with the geographical spread of supply sources and weather and season induced supply variations. In addition, the supply chain is challenged by complex logistics and inventory management aspects, as well as variety of uncertain factors.

Almost every step in the supply chain may have uncertainty factors. Some uncertainties are inherited from the biomass supply part, others from the energy generation and demand parts. Weather conditions and technical disruptions affect harvesting time windows as well as biomass yield and quality. Transportation and logistic uncertainties involve fleet availability, storage and road conditions, all of which induce unpredictable supply. Technological innovation and government policies and incentives greatly affect the competitiveness of investments and operations.

Researchers have looked at biomass-bioenergy markets in settings varying from a single feedstock and a single consumer to integrated settings in the local economy or other energy markets. Despite all of the challenges and uncertain factors, we are aware of only two multi-stage stochastic optimization models which aim to cover the entire supply chain. Cundiff, Dias, and Sherali (1997) develop a multi-stage linear program that considered the impact of weather conditions during the growth season and harvesting period. They include four scenarios, which take into consideration good and bad weather in each period. They considered one type of biomass and allowed storage capacity expansions with fixed locations, but no discrete investment decisions about new capacities or locations were allowed. Walther, Schatka, and Spengler (2012) present a strategic stochastic optimization model for investments related to production networks for synthetic bio-diesel in North-west Germany. Neither of these models include active drying, or the terminal concept which we include in the model presented in this chapter.

Van Tilburg, Egging, and Londo (2006) present the BIOTRANS model, a general model for the biomass to biofuel supply chain. This model is deterministic, has a cost-minimization focus, and does not consider seasonality, storage, or biomass quality variations, although yields are location specific.

Various papers discuss challenges and recommended future research directions. Sharma, Ingalls, Jones, and Khanchi (2013) indicate that the combined complexity and uncertainty plus non-financial objectives require advanced multi-objective approaches. Yue, You, and Snyder (2014) note that models covering uncertainty in biomass quality as well as energy prices and correlation between uncertain parameters are not handled by any known optimization model.

To a large extent the modeling approaches reported in the literature are inspired by a specific case, resulting in many specialized optimization models for certain parts of the value chain or for specific value chains. Various papers argue for a holistic, integrated approach (e.g., De Meyer et al. (2014), Mafakheri and Nasiri (2013), Shabani, Akhtari, and Sowlati (2013), and Wolfsmayr and Rauch (2014)), taking into account interrelationships, interdependences, and coordination needs between all stakeholders in the supply chain, rather than a single agent.

For more details on the bioenergy supply chain, refer to De Meyer et al. (2014), Mafakheri and Nasiri (2013), Scott, Ho, and Dey (2012), Shabani et al. (2013), Sharma et al. (2013), Wolfsmayr and Rauch (2014), and Yue et al. (2014).

Considering the list of challenges and lack of generalized stochastic models in the literature, we have developed a general framework for optimizing the value chain for bioenergy production under uncertainty. This generic stochastic bioenergy optimization model can be used in many types of analysis, independent of technologies considered, types of biomass used, the user operating in a specific part of the value chain, or the geographical region. Our model is flexible in taking a strategic and/or tactical planning perspective, and considers uncertainty in virtually all aspects relevant to the biomass bioenergy supply chain (although in this chapter we do not cover the model extensions needed to handle mandatory crop rotation and perennial crops). Through an adequate composition of the stochastic scenario tree, even correlations between uncertain parameters could be captured. We illustrate the broad applicability of our approach with two case studies that are entirely different in nature; these studies are presented at the end of the chapter.

The rest of the chapter is organized as follows: in the next section (Section 2), we present the core model, i.e. the minimal set of constraints needed for a functional deterministic model. Since uncertainty is such an important characteristic, Section 3 deals with reformulation of the model into a stochastic one. In Section 4, we extend the core model with additional features. Finally, in Section 5, we present two test cases, one dealing with forest biomass in Norway and the other with miscanthus in France.

2 The core model

The model is based on a network representation of the value chain from production to consumption. The nodes of the network represent activities and processes that the products can undergo. Currently, the model has nodes for production, transformation, storage, and consumption. Arcs between the nodes are used to model flow of commodities or equipment between nodes.

This structure means that the model naturally decomposes into several parts: one for each node type and one for the flows between them. The actual network structure is then provided by data, while the model itself is case-independent. The result is a flexible model, capable of handling complex supply chains, such as the one shown in Figure 1. There, we can see that 'chipping' and 'pelletizing' are representing by the same node type, 'transformation'. This means that both the input and output products have to be

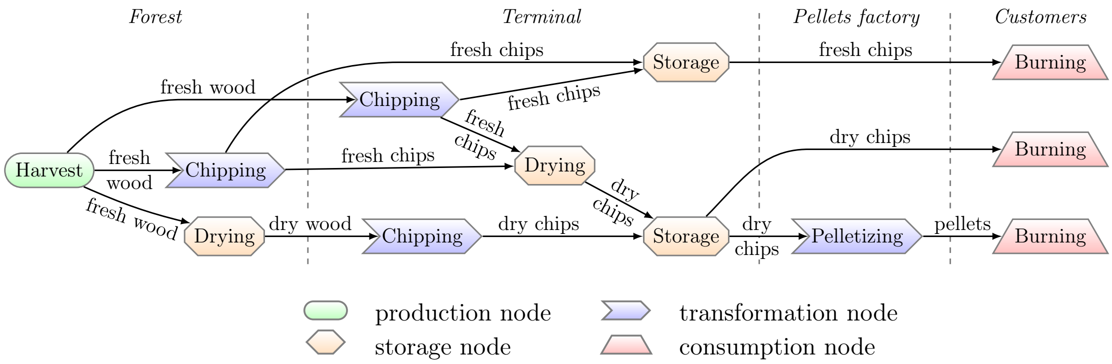


Figure 1: Structure of a stylized forest-biomass value chain.

specified as data. We could save the amount of required data by having separate node types for the two processes, but that would increase the model complexity dramatically; not only would we have to model more node types, but we would also have to model more possible links between them.

2.1 Notation

2.1.1 Indices, sets, and subsets

The model supports the flow of multiple products and their transformation to other products. Each actual product is specified by its product type (referred to simply as 'product') and by the crop that it is made from. Hence, birch chips are a product 'chips' made of crop 'birch'.

Name	Description	Information
$a \in \mathcal{A}$	Transportation arcs	
$c \in \mathcal{C}$	All crops	
$d \in \mathcal{D}$	Dimensions	
$i \in \mathcal{I}$	Production limits	
$j \in \mathcal{J}$	Alternatives	
$n \in \mathcal{N}$	Nodes in production network	
$p \in \mathcal{P}$	All products: input, intermediate and final	
$t \in \mathcal{T}$	Time periods	$\mathcal{T} = 1, \dots, T$
$n \in \mathcal{N}^{\mathrm{P}}$	Production nodes	$\mathcal{N}^{\mathrm{P}} \subset \mathcal{P}$
$n \in \mathcal{N}^{\mathrm{T}}$	Transformation nodes	$\mathcal{N}^{\mathrm{T}} \subset \mathcal{P}$
$n \in \mathcal{N}^{\mathrm{S}}$	Storage nodes	$\mathcal{N}^{\mathrm{S}} \subset \mathcal{P}$
$n \in \mathcal{N}^{\mathrm{C}}$	Consumption nodes	$\mathcal{N}^{\mathrm{C}} \subset \mathcal{P}$
$p \in \mathcal{P}^{\mathrm{B}}$	All basic biomass (untreated, possibly harvested) crops	$\mathcal{P}^{\mathrm{B}} \subset \mathcal{P}$
$p \in \mathcal{P}^{\mathrm{I}}$	All intermediate products (treated, possibly ready for use)	$\mathcal{P}^{\mathrm{I}} \subset \mathcal{P}$
$p \in \mathcal{P}^{\mathrm{F}}$	All final products (ready for use)	$\mathcal{P}^{\mathrm{F}} \subset \mathcal{P}$

Note that $\mathcal{P}^{\mathrm{B}}$, $\mathcal{P}^{\mathrm{I}}$ and $\mathcal{P}^{\mathrm{F}}$ may overlap partially, while sets $\mathcal{N}^{\mathrm{P}}$, $\mathcal{N}^{\mathrm{T}}$, $\mathcal{N}^{\mathrm{S}}$, and $\mathcal{N}^{\mathrm{C}}$ must form a partition of $\mathcal{N}$.

2.1.2 Constants

The flow of products can be measured in a number of dimensions; in bioenergy models, a typical unit is the energy content [GJ] or dry mass content [t^{dm}], while in other applications, a volume measure is more appropriate, e.g., [lm^3] (loose m^3 , e.g., for wood chips). In the model, subscript d in q_d denotes the unit of measurement. Furthermore, each product p has its default dimension d^*_p along which it is measured; it follows that the default units used for the product are $\mathrm{q}_{\mathrm{d}^*_p}$, which we denote by q^*_p. This can then be converted into other units using the unit-conversion parameters $U_{c,p,d}$.

Name	Description	Unit
$C^{\mathrm{A}}_{n_1,n_2,p,d}$	Transport costs	€/q_d
C^{N}_{n}	Operating costs of node n, per period	€
$C^{\mathrm{P}}_{n,c}$	Production costs, per area	€/ha
$C^{\mathrm{Q}}_{n,c}$	Production costs, per volume	€/$\mathrm{q}^*_{\mathrm{p}^{\mathrm{b}}_c}$
$C^{\mathrm{R}^-}_{n,d}$	Transformation costs (per input unit)	€/q_d
$D^{\min}_{n,t,d}$	Demand minimum	q_d
$D^{\max}_{n,t,d}$	Demand maximum	q_d
$E^{\mathrm{D}}_{n,p}$	Efficiency of product p in satisfying demand at node n	
$P_{n,t,d}$	Market price	€/q_d
$\overline{Q}^{\mathrm{A}}_{n_1,n_2,p,d}$	Transportation capacity	q_d
$\overline{Q}^{\mathrm{R}^-}_{n,d}$	Input capacity of a transformation	q_d
$\overline{Q}^{\mathrm{R}^+}_{n,d}$	Output capacity of a transformation	q_d
$Q^{\mathrm{P}}_{n,c,p,t}$	Actual produced volumes at each period	q^*_p
$\overline{Q}^{\mathrm{P}}_{n,c}$	Production capacity (area)	ha
$\mathrm{Q}^{\min}_i$	Minimum production of production limit i	d_i
$\mathrm{Q}^{\max}_i$	Maximum production of production limit i	d_i
$S_{n,d}$	Storage capacity	q_d
$U_{c,p,d}$	Unit conversion ('density')	$\mathrm{q}_d/\mathrm{q}^*_p$
$Y^{\mathrm{R}}_{n,p}$	Transformation efficiency (output per input)	
$Y^{\mathrm{P}}_{n,c}$	Production yield	$\mathrm{q}^*_{\mathrm{p}^{\mathrm{b}}_c}$/ha

2.1.3 Variables

The model has one set of binary decision variables for the node usage, used for calculating operating costs, plus several continuous, non-negative variables for modeling the flows and volumes.

Name	Description	Unit
$f_{n_1,n_2,c,p,t}$	Transportation flow	q^*_p
$q^{\mathrm{P}}_{n,c,p,t}$	Production quantity	q^*_p
$q^{\mathrm{D}}_{n,c,p,t}$	Consumption quantity	q^*_p
$r^{\mathrm{in}}_{n,c,p,t}$	Transformation input quantity	q^*_p
$r^{\mathrm{out}}_{n,c,p,t}$	Transformation output quantity	q^*_p
$s_{n,c,p,t}$	Storage levels at the end of a period	q^*_p
$z_{n,t}$	Whether a node is used in a period	0/1

Note that we model flows using continuous variables, ignoring the fact that they are carried by (discrete) vehicles. This is a natural simplification for a tactical/strategic model, with periods spanning weeks or even months.

2.2 Production nodes

Production nodes are the sources of the biomass, i.e. the fields or forests. Each node can produce multiple crops and products, because one field or forest can be planted with more than one crop and some crops can be harvested in several ways. Note that the latter effect could also be achieved by letting the production node only 'produce' the plant, and by modeling the harvesting process in separate transformation nodes. This would give us more control over the harvesting techniques and allow for associating different efficiencies and costs to the harvesting methods—assuming that we (the optimizing agent) actually can control these processes.

It follows that the constraints describing the production nodes are necessarily case-dependent. If we, for example, do not have any control over the production, the produced volumes are input data to the model:

$$q^{\mathrm{P}}_{n,c,\mathrm{p}^{\mathrm{b}}_c,t} = Q^{\mathrm{P}}_{n,c,\mathrm{p}^{\mathrm{b}}_c,t}\,,$$

where $\mathrm{p}^{\mathrm{b}}_c \in \mathcal{P}^{\mathrm{B}}$ denotes the product of harvesting crop c.

If we, on the other hand, control the production, then we have to distinguish between cases where there is a harvest time when we harvest all of the crop (as with grains) and cases where the crops grow continuously and we harvest only a part of the crop (as with forests). In the former situation, production is limited by the planted area $\overline{Q}^{\mathrm{P}}_{n,c}$ and the yields $Y^{\mathrm{P}}_{n,c}$:

$$\sum_{t\in\mathcal{T}} q^{\mathrm{P}}_{n,c,\mathrm{p}^{\mathrm{b}}_c,t} \le \overline{Q}^{\mathrm{P}}_{n,c}\, Y^{\mathrm{P}}_{n,c}\,, \qquad n \in \mathcal{N}^{\mathrm{P}}, \tag{1}$$

where the unspecified indices run over their default sets, i.e. $c \in \mathcal{C}$. We use this simplification throughout the paper.

In the latter case with partial harvesting, we specify the production limits for each product in some given time period, to ensure sustainability. Each production limit $i \in \mathcal{I}$ is specified for every production node (n_i), crop (c_i), and dimension (d_i) and time interval $[\mathrm{t}^{\mathrm{S}}_i, \mathrm{t}^{\mathrm{E}}_i]$ at which it was measured:

$$\mathrm{Q}^{\min}_i \le \sum_{\mathrm{t}^{\mathrm{S}}_i \le t \le \mathrm{t}^{\mathrm{E}}_i} U_{c,p,d}\; q^{\mathrm{P}}_{\mathrm{n}_i,\mathrm{c}_i,\mathrm{p}^{\mathrm{b}}_{\mathrm{c}_i},t} \le \mathrm{Q}^{\max}_i \qquad i \in \mathcal{I}\,. \tag{2}$$

If a harvested crop results in several products, as it does for trees, we can model this through a transformation node, attached to the production node.

Note that the planted area is taken as input data, i.e., the model does not control what gets planted where in the first model version. This is because in all of our test cases, we optimized other parts of the supply chain and, therefore, could not make these decisions. However, the model is easily extensible to situations where the planting/sowing is a part of the decision process.

2.3 Transformation nodes

Transformation nodes convert products; they represent processes like chipping of trees, production of pellets from chips, and gasification of biomass.

There are many possible types of transformation nodes, depending on the number of input and output products, and by the way they combine. In our model, we use two different types. First, we have a one-to-many transformation node n, in which an input product $\mathrm{p}_n^{\mathrm{in}}$ gets transformed into one or more output products. This is used to model, for example, tree harvesting, where we separate output products logs and branches.

The produced volumes of each output product p are controlled using the transformation yield per unit of input product. In addition, each transformation node has a conversion capacity in terms of the input volume, possibly along several dimensions (weight, volume).

$$r^{\mathrm{out}}_{n,c,p,t} = Y^{\mathrm{R}}_{n,\mathrm{p}_n^{\mathrm{in}}} \cdot r^{\mathrm{in}}_{n,c,\mathrm{p}_n^{\mathrm{in}},t} \qquad n \in \mathcal{N}^{\mathrm{T}} \tag{3}$$

$$\sum_{c\in\mathcal{C}} U_{c,\mathrm{p}_n^{\mathrm{in}},d}\; r^{\mathrm{in}}_{n,c,\mathrm{p}_n^{\mathrm{in}},t} \leq z_{n,t} \cdot \overline{Q}^{\mathrm{R}^-}_{n,d} \qquad n \in \mathcal{N}^{\mathrm{T}} \tag{4}$$

The other type of transformation nodes in our model represents many-to-one transformations, where the input products are alternatives, i.e., the output product can be made using any of the inputs or their mixture. For example, in the miscanthus supply chain that we present in Section 5.2, pellets can be made from both bales and baled chips.[1]

The constraints are similar to the ones listed above, except that we have one output product $\mathrm{p}_n^{\mathrm{out}}$ and the conversion capacity is given in terms of the output volumes.

$$r^{\mathrm{out}}_{n,c,\mathrm{p}_n^{\mathrm{out}},t} = \sum_{p\in\mathcal{P}} Y^{\mathrm{R}}_{n,p} \cdot r^{\mathrm{in}}_{n,c,p,t} \qquad n \in \mathcal{N}^{\mathrm{T}} \tag{3'}$$

$$\sum_{c\in\mathcal{C}} U_{c,\mathrm{p}_n^{\mathrm{out}},d}\; r^{\mathrm{out}}_{n,c,\mathrm{p}_n^{\mathrm{out}},t} \leq z_{n,t} \cdot \overline{Q}^{\mathrm{R}^+}_{n,d} \qquad n \in \mathcal{N}^{\mathrm{T}} \tag{4'}$$

Another example of this type of transformation is gasification of biomass, though this would require removing the crop-subscript c from the output product in (3'):

$$r^{\mathrm{out}}_{n,\mathrm{p}_n^{\mathrm{out}},t} = \sum_{p\in\mathcal{P},c\in\mathcal{C}} Y^{\mathrm{R}}_{n,p} \cdot r^{\mathrm{in}}_{n,c,p,t} \qquad n \in \mathcal{N}^{\mathrm{T}}$$

Note that having one many-to-one transformation node is very similar to having one transformation node for each input product p, with a one-to-one transformation from p to $\mathrm{p}_n^{\mathrm{out}}$. The difference is that this would imply separate capacity for each input product, and to represent a shared input capacity additional transformation nodes would have to be introduced.

For both presented transformation node types, it is natural to attach costs to the incoming products. In the former, there is only one product that gets split into several different ones, while in the latter, the output can be produced from several inputs, where each one can have different transformation costs.

[1] Note that this transformation is different from another common many-to-one process, namely the assembly of parts into one product. There, all inputs are needed to make the output, typically with fixed proportions—something we have not encountered in any biomass chain that we have studied.

2.4 Storage nodes

A storage node models the storage of one or several products between time periods. Each such node is specified by its capacity limits, which can be given along several dimensions (volume, weight). Storage levels for each product are tracked throughout all time periods:

$$\sum_{c \in \mathcal{C}} \sum_{\substack{p \in \mathcal{P}: \\ \exists s_{n,c,p,t}}} U_{c,p,d} s_{n,c,p,t} \leq z_{n,t} \cdot S_{n,d} \qquad n \in \mathcal{N}^{\mathrm{S}},\ d : \exists\, S_{n,d} \tag{5}$$

$$s_{n,c,p,t} = \gamma_n \cdot s_{n,c,p,t-1} - \sum_{\substack{n' \in \mathcal{N}: \\ (n,n') \in \mathcal{A}}} f_{n,n',c,p,t} + \sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p,t} \qquad n \in \mathcal{N}^{\mathrm{S}},\ c,p,t : \exists\, s_{n,c,p,t}, \tag{6}$$

where $1 - \gamma_n$ represents the mass loss due to storage for one period at node n; γ_n is typically a number very close to one. Note that this model assumes that if the storage can handle several products, then they all share the whole capacity. If each product has its own dedicated part of the storage place, each should be modeled by their own storage node, all placed in the same location.

Initial storage levels can be provided or the model can use cyclic behavior where the storage level in the last period is used as the initial level. The first is modeled as $s_{n,c,p,0} = S_{n,c,p}$ and the second one is modeled as $s_{n,c,p,0} = \gamma_n \cdot s_{n,c,p,T}$. In the second case, the initial storage level is a decision variable whose optimal value becomes an important part of the solution.

2.5 Consumption nodes

Consumption nodes are nodes consuming some of the products. They are specified by a minimum and maximum demand for each period:

$$z_{n,t} \cdot D^{\min}_{n,t,d} \leq \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}^{\mathrm{F}}} E^{\mathrm{D}}_{n,p}\, U_{c,p,d}\, q^{\mathrm{D}}_{n,c,p,t} \leq z_{n,t} \cdot D^{\max}_{n,t,d} \qquad n \in \mathcal{N}^{\mathrm{C}}, \tag{7}$$

where each inequality is created only if the associated parameter exists. Note that if $D^{\min}_{n,t,d} > 0$, then we can fix $z_{n,t} = 1$. Parameters $E^{\mathrm{D}}_{n,p}$ can be used to limit which products can satisfy the demand at each given node.

Furthermore, the customers may have limits on different crop shares in the total amount that they buy. This is modeled similarly to the production limits, i.e., we have a set $\mathcal{K}$ of limits, each specified by its node n^{K}_k and a minimum and/or maximum proportion of crop $\mathrm{c}^{\mathrm{K}}_k$ in the mix, denoted by $k^{\min}_k$ and $k^{\max}_k$. The constraints are then

$$k^{\min}_k \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}^{\mathrm{F}}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,c,p,t} \leq \sum_{p \in \mathcal{P}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,\mathrm{c}^{\mathrm{K}}_k,p,t} \leq k^{\max}_k \sum_{c \in \mathcal{C}} \sum_{p \in \mathcal{P}^{\mathrm{F}}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,c,p,t} \qquad k \in \mathcal{K}, \tag{8}$$

where each inequality is created only if the corresponding limit $k^{\min}_k$ or $k^{\max}_k$ exists for given k.

2.6 Transportation and flows

Transport capacity can be provided for any measurement unit:

$$\sum_{c \in \mathcal{C}} U_{c,p,d}\, f_{n_1,n_2,c,p,t} \le \overline{Q}^{\mathrm{A}}_{n_1,n_2,p,d} \qquad (n_1,n_2) \in \mathcal{A} \tag{9}$$

Moreover, in all nodes except storage nodes, we need conservation-of-flow constraint, that is, the total consumption, outward transportation flows, and transformation inputs must equal production, inward transportation flows, and transformation output, in every time period:

$$q^{\mathrm{P}}_{n,c,p,t} + \sum_{n' \in \mathcal{N}} f_{n',n,c,p,t} + r^{\mathrm{out}}_{n,c,p,t} = q^{\mathrm{D}}_{n,c,p,t} + \sum_{n' \in \mathcal{N}} f_{n,n',c,p,t} + r^{\mathrm{in}}_{n,c,p,t}\,, \tag{10}$$

for all $n \in \mathcal{N} \setminus \mathcal{N}^{\mathrm{S}}$. In the model implementation, only the required variables will actually be created at the different types of nodes: production nodes will only have production and outward transportation flows, etc.

2.7 Objective function

The exact form of the objective function is case dependent: if there is a given demand that has to be satisfied, then it is natural to minimize the costs of doing so. If, on the other hand, we can freely choose how much to deliver to each customer, then one also has to take into account income and maximize the overall profit instead.

Therefore, instead of stating the complete objective function, we list its components: the income and different types of costs.

income from sale
$$\sum_{n,c,p,t,d} U_{c,p,d}\, P_{n,t,d}\, q^{\mathrm{D}}_{n,c,p,t} \tag{11a}$$

production/harvesting costs,
using constraint (1)
$$\sum_{n,c,p,t} \left(\frac{C^{\mathrm{P}}_{n,c}}{Y^{\mathrm{P}}_{n,c}} + C^{\mathrm{Q}}_{n,c} \right) q^{\mathrm{P}}_{n,c,p,t} \tag{11b}$$

transportation costs
$$\sum_{n_1,n_2,c,p,t,d} C^{\mathrm{A}}_{n_1,n_2,p,d}\, U_{c,p,d}\, f_{n_1,n_2,c,p,t} \tag{11c}$$

transformation costs
$$\sum_{n,c,p,t,d} C^{\mathrm{R}-}_{n,d}\, U_{c,p,d}\, r^{\mathrm{in}}_{n,c,p,t} \tag{11d}$$

node-usage costs
$$\sum_{n,t} C^{\mathrm{N}}_{n}\, z_{n,t} \tag{11e}$$

Note that the node-usage variables $z_{n,t}$ are really needed only for nodes with non-zero costs C^{N}_n; without these costs, we could omit the variables, though this would require slight reformulation of some of the constraints. We discuss this further in Section **??**.

3 Stochastic formulation

In most bioenergy applications, the problems are subject to multiple uncertainties, as future demand, supply (yield), and prices are seldom known with precision ahead of time. It is therefore natural to use a model that can handle at least some of the uncertainties. There are several approaches to optimization under uncertainty, of which we have chosen stochastic programming, in particular its formulation with stochastic variables represented by discrete scenarios.

It is relatively easy to convert a deterministic model to a two-stage stochastic model since all we need to add is a scenario index to all the stochastic entities. However, we want a general multi-stage formulation, because this gives us the freedom to change the complexity of the stochastic representation solely using data, i.e., without changing the model.

There are two basic approaches to converting a deterministic model into a general stochastic one, illustrated graphically in Table 1.

In a scenario-based formulation, we simply add a scenario index to all time-dependent model entities and then enforce the scenario-tree structure using so-called non-anticipativity constraints (Birge and Louveaux, 1997). While adding the extra index is easy, adding the constraints is not a trivial task. In addition, it will make the model bigger. Even in the very simple scenario tree in Table 1, we have five groups of constraints, one with four nodes and four with two nodes. To enforce the equality, one typically sets all nodes equal to the first one, resulting in $k-1$ constraints for a group of k nodes. In our case, this means $3 + 4 \times 1 = 7$ constraints ... for every stochastic variable.

Alternatively, in a node-based approach, we replace the time index by a scenario-tree-node index for all time-dependent model entities. The structure of the scenario tree is then described by providing a parent (or *predecessor*) node $\mathrm{Pa}(v)$ to each node $v \in \mathcal{V}$: $\mathrm{Pa}(1) = 0, \mathrm{Pa}(2) = 1, \ldots, \mathrm{Pa}(11) = 8, \mathrm{Pa}(12) = 11$. While this avoids extra constraints, it makes the model more difficult to read, especially if we have constraints that cover more than two periods (as we can see in Table 1).

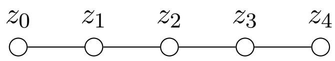
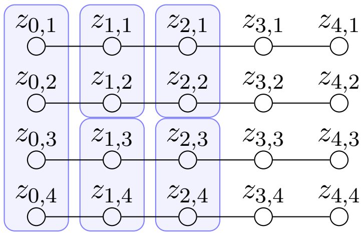
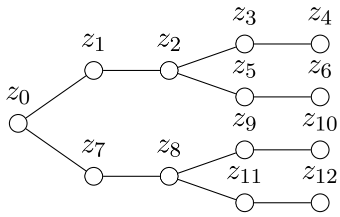
Hence, we use a 'hybrid' approach where we use the node-based approach for all model entities, but provide extra data structures that allow us to use scenario numbers in the model formulation where needed. In particular, we define $\mathrm{SP}(t, s)$ to specify the node in period $t \in \mathcal{T}$ of scenario $s \in \mathcal{S}$; taking an example from the last figure of Table 1, the scenario-tree node corresponding to $t = 2$ in scenario $s = 3$ is z_8, so we have $\mathrm{SP}(2, 3) = 8$. Using this parameter, we can write the constraint from Table 1 as

$$\sum_{t=0}^{4} z_{\mathrm{SP}(t,s)} = 1 \,, \quad s \in \{1, \ldots, 4\}$$

This way, we have eliminated the major disadvantage of the node-based formulation, without the extra constraints needed for the scenario-based approach.

With this notation in place, rewriting the deterministic model from Section 2 into a stochastic one becomes relatively straightforward. For all of the variables, we replace the time index t with the new stochastic-node index $v \in \mathcal{V}$. For links between periods, $t-1$

Table 1: Comparison of the two types of stochastic model representations, on a five-period tree with binary branching in periods 2 and 4. The blue boxes in the scenario-based formulation denote groups of nodes that have to be set equal using non-anticipativity constraints.

<table>
<tr><th>model type</th><th>graphical representation</th><th>example constraint</th></tr>
<tr><td>deterministic</td><td>
</td><td>$\sum_{t=0}^{4} z_t = 1$</td></tr>
<tr><td>stochastic
scenario-based</td><td>
</td><td>$\sum_{t=0}^{4} z_{t,s} = 1,$
$\forall s \in \{1, \dots, 4\}$</td></tr>
<tr><td>stochastic
node-based</td><td>
</td><td>$z_v + z_{\mathrm{Pa}(v)} + z_{\mathrm{Pa}(\mathrm{Pa}(v))} + z_{\mathrm{Pa}(\mathrm{Pa}(\mathrm{Pa}(v)))} + z_{\mathrm{Pa}(\mathrm{Pa}(\mathrm{Pa}(\mathrm{Pa}(v))))} = 1,$
$\forall v \in \{4, 6, 10, 12\}$</td></tr>
</table>

becomes $\mathrm{Pa}(v)$, while links further back in time are easiest written using the $\mathrm{SP}(t,s)$ table.

As we have discussed in Section 2, one can construct several different objective functions using the elements presented in (11), depending on the optimizing agent. Here, we present a stochastic version of one of the variants, namely an agent maximizing the overall profit in the supply chain. For simplicity, we assume a risk-neutral agent and therefore maximize the expected profit, i.e., the probability-weighted profit over all scenario nodes.

The whole model from Section 2, with equation numbers referring to the original deterministic ones, is then as follows:

$$\begin{aligned} \text{maximize} \sum_{v \in \mathcal{V}} \Pr(v) \Bigg(& \sum_{n,c,p,d} U_{c,p,d}\, P_{n,\mathrm{Per}(,)v} d\, q^{\mathrm{D}}_{n,c,p,v} - \sum_{n,c,p} \left(\tfrac{C^{\mathrm{P}}_{n,c}}{Y^{\mathrm{P}}_{n,c}} + C^{\mathrm{Q}}_{n,c} \right) q^{\mathrm{P}}_{n,c,p,v} \\ & - \sum_{n_1,n_2,c,p,d} C^{\mathrm{A}}_{n_1,n_2,p,d}\, U_{c,p,d}\, f_{n_1,n_2,c,p,v} \\ & - \sum_{n,c,p,d} C^{\mathrm{R}^-}_{n,d}\, U_{c,p,d}\, r^{\mathrm{in}}_{n,c,p,v} - \sum_{n} C^{\mathrm{N}}_{n}\, z_{n,v} \Bigg) \end{aligned} \tag{11$^{\mathrm{s}}$}$$

subject to

$$\sum_{t\in\mathcal{T}} q^{\mathrm{P}}_{n,c,p,\mathrm{SP}(t,s)} \leq \overline{Q}^{\mathrm{P}}_{n,c}\, Y^{\mathrm{P}}_{n,c}\,, \qquad n\in\mathcal{N}^{\mathrm{P}}, p\in\mathcal{P}^{\mathrm{B}} \tag{1$^{\mathrm{s}}$}$$

$$\mathrm{Q}^{\min}_i \leq \sum_{\mathrm{t}^{\mathrm{S}}_i\leq t\leq \mathrm{t}^{\mathrm{E}}_i} U_{c,p,d}\; q^{\mathrm{P}}_{\mathrm{n}_i,\mathrm{c}_i,\mathrm{p}^{\mathrm{b}}_{\mathrm{c}_i},\mathrm{SP}(t,s)} \leq \mathrm{Q}^{\max}_i \qquad i\in\mathcal{I} \tag{2$^{\mathrm{s}}$}$$

$$r^{\mathrm{out}}_{n,c,p,v} = Y^{\mathrm{R}}_{n,\mathrm{p}^{\mathrm{in}}_n}\cdot r^{\mathrm{in}}_{n,c,\mathrm{p}^{\mathrm{in}}_n,v} \qquad n\in\mathcal{N}^{\mathrm{T}} \tag{3$^{\mathrm{s}}$}$$

$$\sum_{c\in\mathcal{C}} U_{c,\mathrm{p}^{\mathrm{in}}_n,d}\; r^{\mathrm{in}}_{n,c,\mathrm{p}^{\mathrm{in}}_n,v} \leq z_{n,v}\cdot \overline{Q}^{\mathrm{R}^-}_{n,d} \qquad n\in\mathcal{N}^{\mathrm{T}} \tag{4$^{\mathrm{s}}$}$$

$$\sum_{c\in\mathcal{C}}\; \sum_{\substack{p\in\mathcal{P}:\\ \exists s_{n,c,p,v}}} U_{c,p,d}\; s_{n,c,p,v} \leq z_{n,v}\cdot S_{n,d} \qquad n\in\mathcal{N}^{\mathrm{S}} \tag{5$^{\mathrm{s}}$}$$

$$\begin{aligned} s_{n,c,p,v} = {} & \gamma_n\cdot s_{n,c,p,\mathrm{Pa}(v)} \\ & - \sum_{\substack{n'\in\mathcal{N}:\\ (n,n')\in\mathcal{A}}} f_{n,n',c,p,t} + \sum_{\substack{n'\in\mathcal{N}:\\ (n',n)\in\mathcal{A}}} f_{n',n,c,p,t} \end{aligned} \qquad n\in\mathcal{N}^{\mathrm{S}} \tag{6$^{\mathrm{s}}$}$$

$$z_{n,v}\cdot D^{\min}_{n,t,d} \leq \sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}^{\mathrm{F}}} E^{\mathrm{D}}_{n,p}\, U_{c,p,d}\; q^{\mathrm{D}}_{n,c,p,v} \leq z_{n,v}\cdot D^{\max}_{n,t,d} \qquad n\in\mathcal{N}^{\mathrm{C}} \tag{7$^{\mathrm{s}}$}$$

$$k^{\min}_k \sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}^{\mathrm{F}}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,c,p,v} \leq \sum_{p\in\mathcal{P}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,\mathrm{c}^{\mathrm{K}}_k,p,v} \leq k^{\max}_k \sum_{c\in\mathcal{C}}\sum_{p\in\mathcal{P}^{\mathrm{F}}} q^{\mathrm{D}}_{n^{\mathrm{K}}_k,c,p,v} \qquad k\in\mathcal{K} \tag{8$^{\mathrm{s}}$}$$

$$\sum_{c\in\mathcal{C}} U_{c,p,d}\; f_{n_1,n_2,c,p,v} \leq \overline{Q}^{\mathrm{A}}_{n_1,n_2,p,d} \qquad (n_1,n_2)\in\mathcal{A} \tag{9$^{\mathrm{s}}$}$$

$$q^{\mathrm{P}}_{n,c,p,v} + \sum_{n'\in\mathcal{N}} f_{n',n,c,p,v} + r^{\mathrm{out}}_{n,c,p,v} = q^{\mathrm{D}}_{n,c,p,v} + \sum_{n'\in\mathcal{N}} f_{n,n',c,p,v} + r^{\mathrm{in}}_{n,c,p,v} \qquad n\in\mathcal{N}\setminus\mathcal{N}^{\mathrm{S}} \tag{10$^{\mathrm{s}}$}$$

Just like in the deterministic model, we have to decide how to deal with the initial storage levels in (6$^{\mathrm{s}}$). If we used the initial-storage constraint ($s_{n,c,p,0} = S_{n,c,p}$) in the deterministic model, then we can use it without change. If, on the other hand, we used the cyclical storage approach ($s_{n,c,p,0} = \gamma_n \cdot s_{n,c,p,T}$), then we have to decide how to interpret this in the stochastic settings. There are several possible approaches to this, each with merits and problems:

$$s_{n,c,p,0} = \sum_{v\in\mathcal{V}_T} \Pr(v)\cdot s_{n,c,p,v} \tag{12a}$$

$$s_{n,c,p,0} = s_{n,c,p,v} \qquad \forall v\in\mathcal{V}_T \tag{12b}$$

$$s_{n,c,p,0} \leq s_{n,c,p,v} \qquad \forall v\in\mathcal{V}_T\,, \tag{12c}$$

where $v = 0$ is the root node of the scenario tree and $\mathcal{V}_T$ is the set of last-period nodes, $\mathcal{V}_T = \{v\in\mathcal{V} : \mathrm{Per}(v) = T\}$.

The first formulation, (12a), ensures that *on average*, we will finish with the same amount in storage at the end as at the start. This is probably the most natural extension of the deterministic version. It can, however, give undesired effects; for example, assume that we include demand as the only stochastic parameter in the model and that the optimal storage level in the first period is $s^*_0 > 0$. Constraint (12a) will then cause the end-of-horizon storage level to be higher than s^*_0 in low-demand scenarios and lower—possibly even zero—in scenario(s) with the highest demand. The end-of-horizon storage level could even be zero in the 'average' scenario, which would imply that the results found by the model would not provide proper guidance for two successive average years (since the average year needs positive initial storage levels, but finishes with empty storage).

This problem is avoided in version (12b), where the final storage has to be equal in all scenarios. Unfortunately, this creates new problems. If we, for example, have one scenario with very low demand, it may be optimal to end up with a high level of storage at the end. With (12b), however, the final storage at the end of the low-demand scenario has to be the same as in high-demand scenarios. The optimal solution might then be to produce lower amounts of biomass, which may imply a huge loss of profit potential. The last issue is resolved in (12c), where we allow the low-demand scenarios to end up with more storage. Another way of resolving the issues in (12b) is to add extra time at the end of the horizon, to allow the models to settle on common storage levels (Thapalia, Wallace, and Kaut, 2009).

In contrast to (12a) that ensures that *on average*, we will finish with enough storage to continue another *average* year, (12c) ensures that this happens in every scenario. In this sense, the latter represents a more conservative risk attitude. For this reason, this is the formulation we use in our implementation.

4 Extensions to the model

In this section, we present several extensions to the model. We formulate them for the stochastic version of the model, but note that they can be used for the deterministic version as well. Each of these adds specific functionality to the model. They can be combined, depending on what is needed for each case.

4.1 Terminals

By *terminal*, we mean an area with multiple facilities (e.g., transformation and storage). In the context of the model, a *terminal* is a grouping of one or more nodes. There are at least two uses for terminals in the model: they allow us to model costs associated with running the terminal in addition to costs for each facility/node. For this, we just need to associate a binary variable y_g to each terminal $g \in \mathcal{G}$ and then require that the nodes belonging to the terminal, $n \in \mathcal{N}^{\mathrm{G}}_g$, can be used only if y_g is equal to one:

$$z_{n,v} \le y_g \qquad\qquad n \in \mathcal{N}^{\mathrm{G}}_g,\ v, g : \exists\, z_{n,v} \text{ and } \exists\, y_g\,.$$

More importantly, in some applications the goal of the model is to establish a new terminal, by choosing one from a list of candidate locations. We model this in a more general way, using a set $\mathcal{J}$ of 'alternatives'. For each alternative $j \in \mathcal{J}$, we then specify a group of terminals, $\mathcal{G}_j$, with associated lower and upper bounds on the number of terminals to be opened, $\underline{G}_j$ and $\overline{G}_j$:

$$\underline{G}_j \le \sum_{g \in \mathcal{G}_j} y_g \le \overline{G}_j \qquad\qquad j \in \mathcal{J}$$

If we are to choose exactly one from all of the potential terminals, we do this using $\mathcal{J} = \{1\}$, $\mathcal{G}_1 = \mathcal{G}$, and $\underline{G}_1 = \overline{G}_1 = 1$. This formulation assumes that the decision regarding terminals is made at the start of the first period and is valid for the entire duration of the model. An alternative would be to add a time index to these decisions, so that they can be postponed. Note that this would increase the number of binary variables, and hence, the solution time.

In addition to the above constraints, we would also need to add the costs of opening the terminals, $\sum_g C_g^{\mathrm{G}}\, y_g$, to the objective function. Note that we could easily add terminal-usage costs (per period) as well, though it would require extra binary variables.

4.2 Drying

In forestry applications, it is important to model drying of the wood. There are at least two approaches to modeling the process in an optimization model, where the difference is whether we treat the moisture content continuously or discretize it. In the former case, we attach a variable for moisture content to each product that needs drying and then track how this value decreases over time. The difficulty with this approach is that one needs to express the energy content of the product as a function of the moisture content—and since this relationship is non-linear, this would mean employing some kind of piecewise-linear approximation (Van Dyken, Bakken, and Skjelbred, 2010).

In the latter approach, we expand the set of products to contain additional information about moisture content. For example, we replace the product type 'log' by 'log 20%,' 'log 40%,' 'log 60%,' and 'log 80%,' each with the appropriate energy content. With this approach, drying is modeled as a change in the product type, after given time in storage.

In our implementations, we have used the latter approach, where the transformation does not depend only on the time spent in the storage, but also on the actual period – the model was developed for a customer in Norway, where wood dries only during the summer.

In the model, the possibility that a product p of crop c can transition into another product in storage node n is signaled by existence of $r = \mathrm{R}_{n,c,p}$. The transition is then specified by its output product $\mathrm{p}_r^{\mathrm{out}}$, the last period in which the product must arrive to the storage in order to undergo the transition $\mathrm{t}_r^{\mathrm{in}}$, the first time period when the transition is finished and the output product can be taken out of the storage $\mathrm{t}_r^{\mathrm{out}}$, and finally, the mass loss during the transition κ_r. In our particular case, we say that all fresh wood that is in storage by the end of April is dry at the start of September, so we have $\mathrm{t}_r^{\mathrm{in}} = 3$ and $\mathrm{t}_r^{\mathrm{out}} = 9$.

With this notation, (6) is modified in the following way:

$$\begin{aligned} s_{n,c,p,v} = \gamma_n \cdot s_{n,c,p,\mathrm{Pa}(v)} - \sum_{\substack{n' \in \mathcal{N}: \\ (n,n') \in \mathcal{A}}} f_{n,n',c,p,v} + \sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p,v} \\ - \gamma_n^{\mathrm{t}_r^{\mathrm{out}} - \mathrm{t}_r^{\mathrm{in}}} \cdot s_{n,c,p,\mathrm{t}_r^{\mathrm{in}}} \quad \text{if exists } r = \mathrm{R}_{n,c,p} \text{ and } \mathrm{Per}(v) = \mathrm{t}^{\mathrm{out}}_{\mathrm{R}_{n,c,p}} \\ + \sum_{\substack{p' \in \mathcal{P}: \\ \exists\, r = \mathrm{R}_{n,c,p'} \\ p = \mathrm{p}_r^{\mathrm{out}}, \mathrm{Per}(v) = \mathrm{t}_r^{\mathrm{out}}}} \gamma_n^{\mathrm{t}_r^{\mathrm{out}} - \mathrm{t}_r^{\mathrm{in}}} \cdot (1 - \kappa_r) \cdot s_{n,c,p',\mathrm{t}_r^{\mathrm{in}}} \end{aligned} \tag{6'}$$

With the values stated above, the second line ensures that in September, the fresh wood will be removed from storage, while the third line replaces it with a corresponding amount of dry wood (minus the losses κ_r). Note that this assumes that no fresh trees are removed from storage during the drying period.[2]

An alternative, and simpler, way of modeling the drying process would be to assume a constant drying rate throughout the year, that is, to assume that crop c stored at node n needs $\Delta t^{\mathrm{d}}_{c,n}$ time periods to transition from product p^{w} to product p^{d}. If we assume that only fresh crops enter storage and only dry crops leave storage, then the storage equation (6) would become

$$\begin{aligned} s_{n,c,p^{\mathrm{w}},v} &= \gamma_n \cdot s_{n,c,p^{\mathrm{w}},\mathrm{Pa}(v)} + \sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p^{\mathrm{w}},v} - \sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p^{\mathrm{w}},\mathrm{Per}(v) - \Delta t^{\mathrm{d}}_{c,n}} \\ s_{n,c,p^{\mathrm{d}},v} &= \gamma_n \cdot s_{n,c,p^{\mathrm{d}},\mathrm{Pa}(v)} - \sum_{\substack{n' \in \mathcal{N}: \\ (n,n') \in \mathcal{A}}} f_{n,n',c,p^{\mathrm{d}},v} + \sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p^{\mathrm{w}},\mathrm{Per}(v) - \Delta t^{\mathrm{d}}_{c,n}} \end{aligned} \tag{6''}$$

Note that this approach can be used to model *active drying*, we just need to set the storage-using costs accordingly and perhaps introduce a volume-dependent cost as well.

4.3 Tracking equipment

The transformation equation (4) assumes that the capacity is given and constant for each transformation node. This, however, ignores the fact that some transformations need extra equipment—which we refer to as a 'transformation device'—to be present in the transformation node. An example of such a requirement is a mobile chipper needed for off-terminal chipping of wood. Since we normally have only a limited amount of these devices, we need to ensure that each is used only in one place at a time.

There are several ways of handling this in the model. If the devices are so mobile that they can be used in several nodes during one period, we can simply add constraints that limit the overall transformation capacity (summed over all transformation nodes) to the capacity of the available devices. Note that this means that we do not control

[2] If this was a problem, one could add additional product types for 'drying wood' that would be forbidden to be taken out of the storage.

the amount of nodes a device is used at during one period and it also ignores relocation costs. On the other hand, the approach does not add any new variables, and therefore, should not increase the solution time of the model. This approach is used, for example, in Gunnarsson, Rönnqvist, and Lundgren (2004).

If, on the other hand, the transformation devices are costly and/or time-consuming to move, we would need a more detailed representation in our model. There, each transformation device $h \in \mathcal{H}$ is specified by its output product $P^{\mathrm{H}}_h \in \mathcal{P}$, set of nodes where it may be needed $\mathcal{N}^{\mathrm{H}}_h$, and transport costs $C^{\mathrm{H}}_{h,n_1,n_2}$ between nodes n_1 and n_2. We then introduce binary variables $w^h_{n,v}$ denoting whether a transformation device h is present in node n in scenario-tree node v, plus additional variables for tracking their movements where: $w^h_{n,n',v}$ is equal to one if the transformation device h is moved from node n to node n' at the end of period of the scenario-tree node v.

$$w^h_{n,v} = w^h_{n,\mathrm{Pa}(v)} - \sum_{n' \in \mathcal{N}^{\mathrm{H}}_h \setminus \{n\}} w^h_{n,n',\mathrm{Pa}(v)} + \sum_{n' \in \mathcal{N}^{\mathrm{H}}_h \setminus \{n\}} w^h_{n',n,\mathrm{Pa}(v)} \qquad h \in \mathcal{H}, n \in \mathcal{N}^{\mathrm{H}}_h \quad (13)$$

$$\sum_{n \in \mathcal{N}^{\mathrm{H}}_h} w^h_{n,1} = 1 \qquad h \in \mathcal{H} \quad (14)$$

$$z_{n,v} \leq \sum_{h \in \mathcal{H} : n \in \mathcal{N}^{\mathrm{H}}_h} w^h_{n,v} \qquad n \in \mathcal{N}^{\mathrm{T}} \quad (15)$$

Note that it is enough to have (14) only for the first period, since the flow-conservation constraints (13) guarantee that they will hold in all periods. Finally, constraints (15) ensure that transformation nodes cannot be used without the required device in place.

Furthermore, the movement-tracking variables $w^h_{n,n',v}$ do not need to be declared as binary in the model, since they will be automatically integer because of constraints (13). Nevertheless, these variables might increase the solution time significantly, so they should be included only if relocation costs of the devices are high enough compared to other costs. If not, we can simply remove these variables and constraints (13) from the model.

5 Illustrative examples

In this section we present two different cases. In one case, we used a deterministic version of the model, while in the other we included uncertainty. The first case is based on the harvesting of trees, while the raw material in the second one is cultivated. The examples are meant to illustrate the flexibility of the model and show benefits from using an optimization based decision support tool. For this reason, to emphasize the illustrative insights, only three future scenarios are considered in the case with uncertainty.

Both examples have been implemented in the Mosel modelling language and solved using FICO$^{\mathrm{TM}}$ Xpress Optimizer, on an dual-core 2.4 GHz machine with 8 GB of RAM.

Figure 2: The area used in the test case in Section 5.1. The heat plants are denoted by '▲', while '■' denotes the possible locations of terminals.

5.1 Wood-based bioenergy supply chain

In this example we consider a supply chain where wood chips are used in heating plants in western Norway. Wood is available in large quantities in the area, but difficult to harvest due to the steep terrain, which translates to high production costs. The wood chips can be produced from different species, each with different energy contents: pine, spruce, and hardwood.

Production of wood chips can be performed with a mobile wood chipper at the felling areas or at the terminals. The difference between these two alternatives is that transportation of wood chips is cheaper than transportation of logs, but the operational costs of a mobile chipper are larger than those of a stationary chipper at a terminal.

The wood can be used directly (as fresh chips) in some heating plants, but, more commonly, drying is needed before use. Drying can take place at felling areas or at the terminal, and before or after chipping, which means that there are a lot of choices to be made in the upstream part of the supply chain. Trees cut in winter/spring can be used the fall, while trees cut in summer/fall cannot be used until the next fall. This means that timing of felling and storage capacity are important factors to consider in order to meet demand.

Terminals are usually a large area where logs and chips can be dried and stored for protection against rain and snow. Storage and drying at the felling sites is more uncertain and the losses are higher due to less controlled conditions. The heating plants have limited capacity for storage of chips and their demand for chips is largely correlated with weather conditions, which results in seasonal variation. Some of the plants are even closed during summer. An illustration of the value chain in this example is given in Figure 3.

In the example case from the Møre region in Norway, we have the perspective of the society of forest owners. A map of locations in the example case can be seen in Figure 2. Note that ferries are needed for transportation between Håhjem and Ørsta, Håhjem and Årø, and Aspøya and Årø, but Malo and Årø are connected by tunnel.

We want to decide timing and area for felling, when and where to chip, and where to deliver. We also want to make decisions regarding if and where a terminal should be opened. In the example, we have three existing heating plants and three candidate locations for a terminal. Another important input parameter is the available wood from different species; these data are generated by using a GIS tool based on data from the Norwegian government. The costs incurred are production costs for felling, transportation costs, chipping costs, and costs related to storage. Sales amounts to heating plants are measured in energy output of delivered chips (not volume). We do not include rental costs for the terminal areas. In negotiations, profit can be used as a guideline for the forest owners to decide how much they are willing to pay in rent for a terminal. The objective function is to maximize total profit.

We want to decide timing and area for felling, when and where to chip, and where to deliver. We also want to make a decision on if and where a terminal should be opened. In the example we have three existing heating plants and three candidate locations for a terminal. Another important input parameter is available wood of different species, this data is generated by the use of a GIS tool based on data from the Norwegian government.

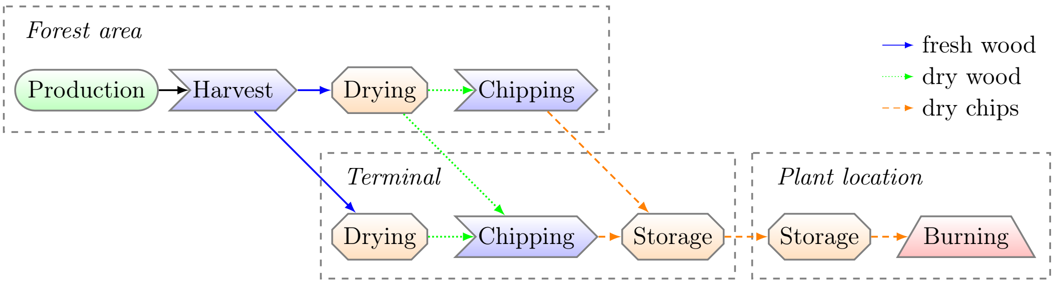


Figure 3: Structure of the forest supply chain in Section 5.1. Note that the case includes three heat plants and three possible terminal locations, each with three forest areas attached. The forest areas are connected only to the nearest terminal, while each terminal can supply all three heat plants.

The costs incurred are production costs for felling, transportation costs, chipping costs, costs related to terminal and costs related to storage. Sales amounts to heating plants are measured in energy output of delivered chips (not volume). The objective function is to maximize total profit.

In this example case, we use all of the extensions presented in Section 4: we have multiple terminals to choose from (4.1), the wood or chips need drying before they can be sent to the heat plants (4.2), and we have one mobile chipper whose location and movement have to be tracked (4.3). For initial storage levels, we use the 'steady-state' formulation with initial levels equal to the end-of-horizon levels, minus some losses.

In addition, we added extra constraints to prohibit harvesting, drying and chipping around a closed terminal. In the core model, the storage constraints (6) allow flow of products through a storage node, even if the node is closed. In our particular case, this was not realistic, so we added the following constraints for the storage nodes at terminals:

$$\sum_{\substack{n' \in \mathcal{N}: \\ (n',n) \in \mathcal{A}}} f_{n',n,c,p,t} \leq \mathrm{M}_{n,c,p,t} \cdot z_{n,t} \,,$$

where the constants $\mathrm{M}_{n,c,p,t}$ are upper bounds on the left-hand side sums, ensuring that the constraints are inactive when $z_{n,t} = 1$. Note that this type of 'big-M constraints' is known to lead to bad LP relaxation, especially with large values of the constants. It is therefore advisable to find as tight upper bounds of the left-hand sides as possible. Note that in our case, it is actually possible to avoid the 'big-M formulation' of these constraints altogether, though at the cost of more changes to the model.

The time perspective of the example case is one year, with each time period being a month in length. The heat plants' demand varies throughout the year. We do not require the demand to be fully satisfied, as there are other sources from which the heat plants can buy their fuel.

Table 2: Summary of results of the case from Section 5.1. All financial values are in thousand NOK.

number of terminals	1	2	3
Revenues	11,020	11,020	11,211
Transp cost	1,572	1,296	1,271
Profit	4,276	4,550	4,642
Total consumption [MWh]	40	40	41
Avg. demand satisfaction	80.6%	80.6%	93.2%
Cost increase for 100% sat.	0.77%	0.19%	0.06%
Value of extra terminal	–	274	92

5.1.1 Numerical results

The example case has been tested imposing how many of the three potential terminal locations should be opened. The differences between opening one, two, and three terminals were then been analyzed and compared. Solution times for the three cases were, respectively, 7, 8, and 31 seconds. Most of the complexity comes from the binary variables used for tracking the mobile chippers—without them, the model solves much quicker.

In the case where we allow only one terminal to be opened, the model uses the one at Aspøya. Looking at the map in Figure 2, this may be surprising, as one could expect the middle terminal (Malo) to be preferred. This, however, can be explained by the topology of the Norwegian west coast. Because of the fjords, the shortest way from Malo to Tingvoll passes by both Årø and Aspøya, and to Ørsta by Årø and Håhjem.

In all the cases, the terminals are used only for storing chips. In other words, all drying and chipping is done in the forest area adjacent to the terminals. This is due to the fact that the disadvantages (higher costs and losses) of forest storage and chipping are more than compensated by lower transportation costs of chips versus whole trees—a finding that is in concordance with Kanzian, Holzleitner, Stampfer, and Ashton (2009).

The numerical results are summarized in Table 2. There, we can see that having two terminals increases the expected profit by 274 thousand NOK, due to decreases in transportation costs. This value should then be compared to the cost of opening the extra terminal. For three terminals, the advantage of the extra terminal decreases to 92 thousand NOK.

While the optimal strategy with one and two open terminals is to satisfy only 80.6% of the energy demand at the three heat plants (on average), the extra costs of satisfying all off the demand are very low. It can therefore be expected that, in practice, one would aim to satisfy all demand. This would decrease the plant operators' incentives to look for alternative fuel sources giving competitors access to the region. For this reason, we assume 100% demand satisfaction in the following figures.

Figure 4 presents the storage levels of the final product (chips) at the terminal and heat plant storages in the case with full demand satisfaction. The graph for the results with one open terminal is on the left, two in the middle, and three to the right. We

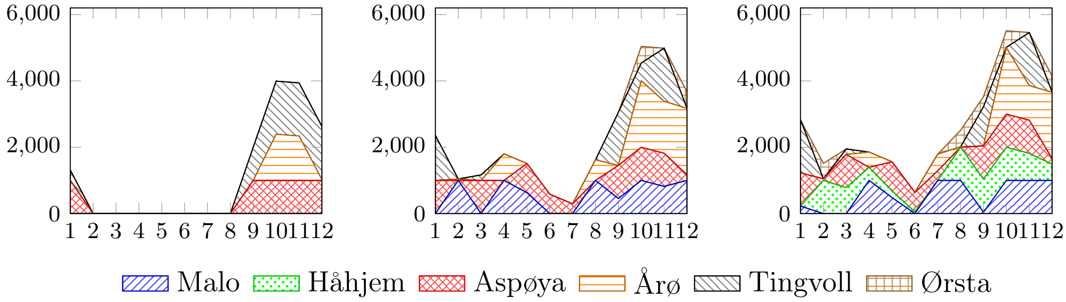


Figure 4: Storage levels of chips at the heat plants and terminals.

can see that there is a big difference between one and two open terminals. Having the extra terminal allows for more storage to build up, which then leads to a decrease in transportation costs, and therefore higher profit. Common to all of the three cases is the buildup of storage during the fall, caused by the fact that the wood has to first dry during the summer.

Finally, Figure 5 presents the flow of chips through storage at the heat plants, for the case of three open terminals and full demand satisfaction. In addition, this figure shows chipping activity at and around the terminals.

We can see that even with all three terminals open, only the heat plant at Tingvoll gets its supply solely from the closest terminal; the Ørsta plant uses two terminals and the Årø plant, being in the middle, gets supplied from all three. This is caused by the fact that we only have one mobile chipper, combined with limited storage volume at both terminals and heat plants.

5.2 Miscanthus transformation plant

Miscanthus is a perennial grass that is increasingly being used for bioenergy purposes. We consider a miscanthus transformation plant located in the Burgundy region of France. The plant previously produced pellets from sugar beets; it now produces pellets from perennial grasses and wood.

Miscanthus is harvested in March and April with three possible options: harvesting and chipping to small chips (2-3 cm) transported directly to the plant, or harvesting and baling with 8–10 cm strands or with 20–30 cm strands. The small chips are compressed, packed, and sold in bags for use in gardening (mulching). The bales can be stored locally at the farmer's location and transported to the plant as needed. All bales should have been picked up by the end of July to make room for autumn crops. The bales with short strands, also called *baled chips*, can be used for pellet production and animal bedding, as well as for energy purposes. Bales with long strands are only used for pellet production. Due to the smaller strands, producing baled chips will incur greater losses both during baling and later handling, but provides flexibility due to multiple sales options. Each harvesting option is modeled as a separate transformation node, see Figure 6. We do not consider drying explicitly in the model as miscanthus is only harvested when moisture

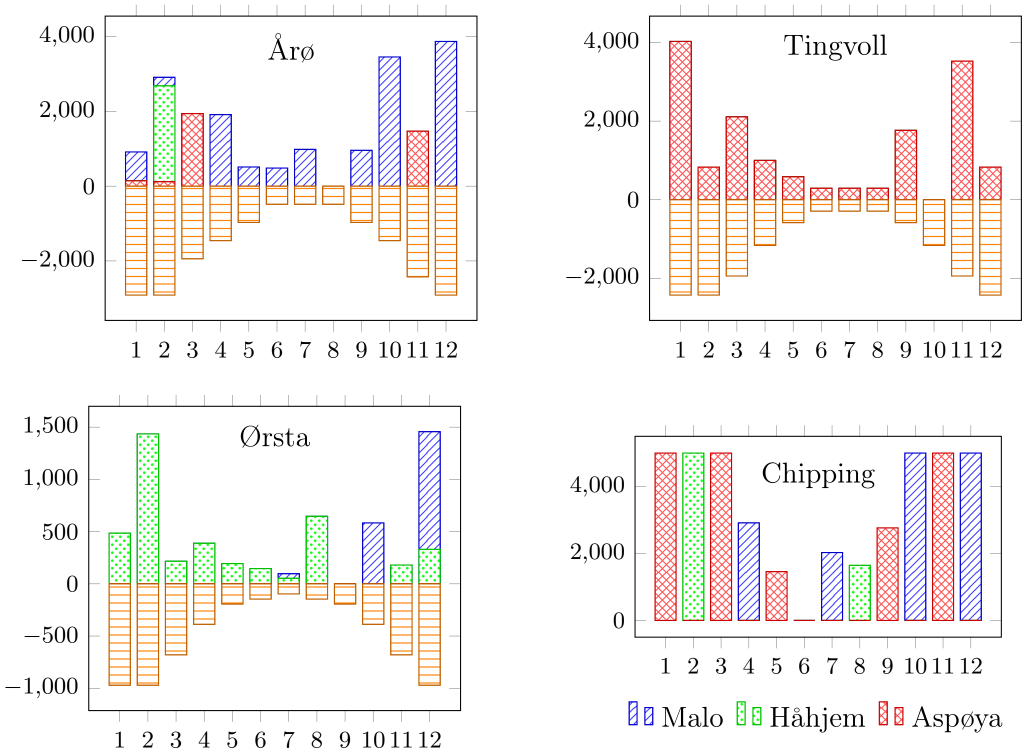


Figure 5: Flow of chips through the storage at the heat plants, plus the chip production, per month. The negative values in the first three charts denote the flow of chips to the plant (i.e. consumption).

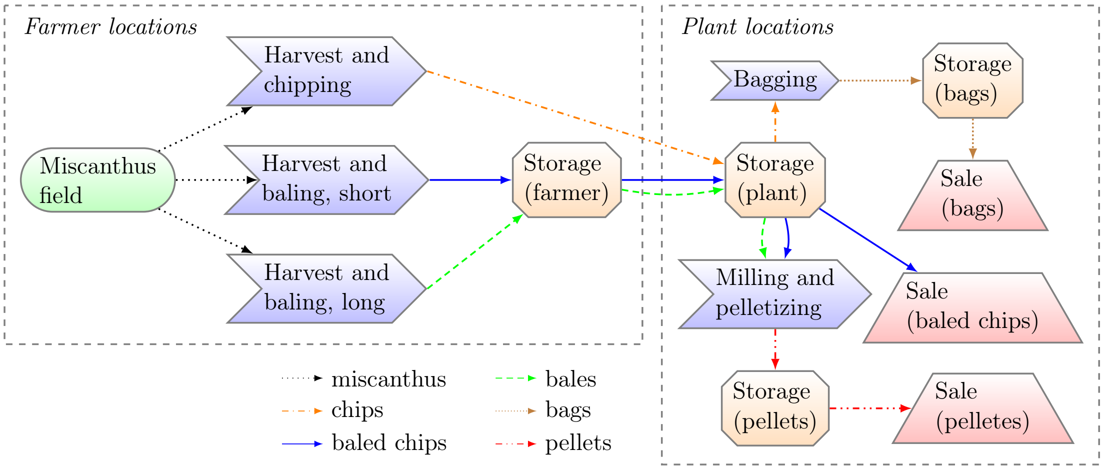


Figure 6: Structure of the flow of miscanthus.

levels are acceptably low.

We consider a case based on current fields planted with miscanthus using estimated yields for mature crops (Rizzo, Martin, and Wohlfahrt, in press 2014) and increased storage capacity compared to the existing state. To illustrate the value of a stochastic model, we consider a situation where the demand of pellets is uncertain. This uncertainty is revealed after harvesting decisions have been made, with demand at the expected level, or 20% higher or lower than that level, with probabilities 0.4, 0.3, and 0.3, respectively. This is an example of a two-stage stochastic programming problem, where the first-stage decisions (how to harvest) are based only on the data available at that time. Table 3 gives a summary of the test case characteristics; note that pellets sell at a slightly higher price than bagged and baled chips. We use a profit-maximizing objective similar to equation (11); we have twelve periods, corresponding to a monthly granularity and we use a cyclic storage behavior.

We start by comparing solutions of the stochastic and deterministic versions of the model, where the latter uses expected values for the stochastic parameter, as presented in the left chart of Figure 7. There, we can see that the results are quite intuitive: the deterministic model produces more bales (long strands) due to smaller losses compared to baled chips (short strands). The stochastic model, on the other hand, produces more baled chips, because these can be used to produce pellets (albeit more expensive than using bales with long strands), and thus compensates for the uncertainty of the pellets' demand.

We proceed by comparing how the two solutions fare in the stochastic environment. To do this, we solve the stochastic model with the first-stage decisions (production variables) fixed to the solution of the deterministic model, and compare the results to the optimal solution of the stochastic model.[3] The results of this comparison are in the right chart of

[3]In our case, the model becomes infeasible because in the low-demand scenarios, we are left with more unsold products than we have storage for. For this reason, we have added additional variables that

Table 3: Test case characteristics

Production:	Producers	60 farmer locations
	Yield	10-18 t^{dm}/ha
	Total area	320 ha
Harvesting:	Chipping	Cost €6.7 / t^{dm}, loss 5%
	Baling, long	Cost €29.2 / t^{dm}, loss 5%
	Baling, short	Cost €29.2 / t^{dm}, loss 10%
Transport:	Distances	Road distances based on OpenStreetMap data
	Cost	€0.4–1.2 / t^{dm}km
Storage:	Capacity	Plant 6720 m^3
	Cost	No storage cost or loss is considered
Sales:	Chips	Max 100 t^{dm}/ month, price €75 / t^{dm}
	Baled chips	Max 200 t^{dm}/ month, price €75 / t^{dm}
	Pellets	Max 300 t^{dm}/ month, price €85 / t^{dm}

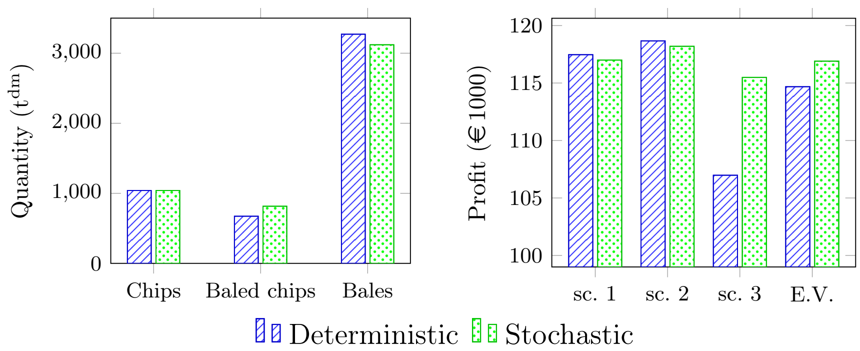


Figure 7: Results of the miscanthus case using the deterministic and stochastic solutions are presented. The left chart displays the transformed quantities of harvested miscanthus. The right chart displays the expected profit of the two solutions when evaluated on the scenario tree.

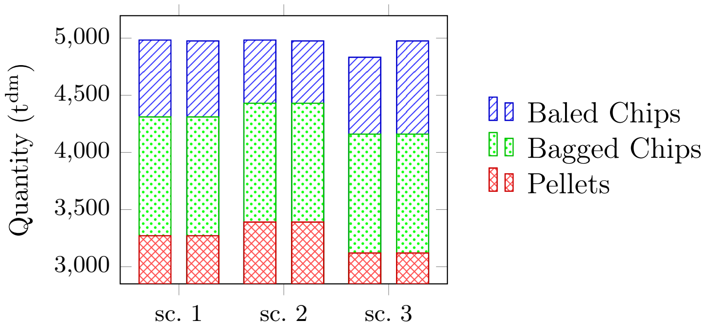


Figure 8: Sold amounts are presented per scenario. The left column in each set represents the deterministic solution; the right column represents the stochastic solution.

Figure 7. We can see that the deterministic solution fares better in the mean-value and high-demand scenario, but suffers in the low-demand scenario, so it is worse on average. Again, this is due to its higher use of the cheaper, yet inflexible, bales. This can be seen explicitly in Figure 8, which shows negligible differences in the sold amounts in the first two scenarios, but higher sales of baled chips for the stochastic solution in the last, low-demand, scenario. Again, this shows that the stochastic solution covers part of the pellets' demand in the first two scenarios using baled chips (causing smaller profits in these two scenarios), which gives it the opportunity to sell more when the demand for pellets is low. In other words, the baled chips are used as a buffer against uncertainty in pellets demand.

Finally, we compute the *value of stochastic solution* (VSS, see Birge and Louveaux, 1997), given as the difference between the expected profit of using the optimal stochastic and deterministic solution. This is the difference between the last two columns shown in the right chart of Figure 7:

$$\mathrm{VSS} = €\,116\,902 - €\,114\,666 = €\,2236\,. \tag{16}$$

This means that in this case, the stochastic solution adds only about 2% to the deterministic one. This is mainly because we have only one stochastic parameter in the model (the demand of pellets). If we let more of the parameters (other demands and prices) be stochastic, the stochasticity would have a higher impact and the value of stochastic solution would increase. On the other hand, it would make interpretation of the results significantly more complicated.

Let us go back to the second value in (16), that is, the expected value of using the expected-value solution (EEV) and compare it to the objective value of the deterministic model, which is €117467. This shows that if we ignore the uncertainty and solve the deterministic model, the reported profit is an over-estimation of the actual profit in the uncertain world (this is a known, and general, observation). It is also interesting to note

allow 'throwing away' products (with neither cost nor income). Obviously, these variables are all zero in the optimal stochastic solution.

Table 4: Results of the case from Section 5.2 for different number of scenarios. Objective values and the VSS are in thousand Euro, while the solution time is in seconds and includes also the time to build the model.

#sc.	obj. value	VSS	#variables	sol. time
1	117.5	–	6 437	0.3
3	116.9	2.2	12 773	0.4
10	116.4	2.0	34 949	0.8
30	116.1	1.8	98 309	2.2
100	116.0	1.8	320 069	7.7
300	116.0	1.8	953 669	28.0
1 000	115.9	1.8	3 171 269	206.7

that the expected profit of the stochastic solution is very close to this figure, showing that its flexibility can almost compensate for the uncertainty.

5.3 Adding more scenarios

So far, we have only tested the stochastic formulations with three scenarios. While this allowed us to examine the difference between scenario solutions, one needs more scenarios to obtain reliable results. To test how many, we have solved the same problem with varying number of scenarios. To keep the model consistent with the three-scenario case, we have used normal distribution with the mean and variance computed from the three scenarios.

The results of the test are presented in Table 4. There, we can see that both the objective value and VSS stabilize at about hundred scenarios. Furthermore, we can see that the case with one thousand scenarios takes three and a half minutes to solve; the time is equally split between building the model and solving it.

Concluding Remarks

We present a new, generic optimization model for strategic and tactical planning of the biomass to bioenergy supply chain under uncertainty. The model structure is flexible and capable of representing relevant characteristics and issues related to the biomass-bioenergy supply chain, including technological process details, capacity limitations in multiple units of measurement, time variability in supply and demand, and uncertainty in virtually all aspects. Two cases of different supply chains illustrate how the model can be parameterized for different types of analysis, and give insight in the effects of uncertainty on optimal decisions.

The model presented can be used by actors in all parts of the supply chain considered. It can improve the decision making processes by giving results and enabling analysis of different possibilities much faster than traditional planning. Integrated in the companies' software this can be a powerful tool for planners and decision makers in an industry with high competition and tight margins.

The flexibility of the model opens for easy expansion and improvement of the model. More details about different stages in the model or tailor-made setups for different companies are examples of likely requests from the industry that can be included. Further development can make the model even more powerful and the value of using the model can be increased, making the bioenergy industry more competitive.

Acknowledgements

This work was partly funded under the EU seventh Framework Programme by the LogistEC project No. 550 311858: Logistics for Energy Crops' Biomass. The views expressed in this work are the sole responsibility of the authors and do not necessary reflect the views of the European Commission. This work was partly funded by Regionalt forskningsfond Midt-Norge through the project 'Fra skog til energi' (ES 217558).

We are grateful to Philippe Béjot (Bourgogne Pellets Cooperative) who kindly provided details on the miscanthus case.

Maps and distance matrices were created using data from OpenStreetMap, © OpenStreetMap contributors.

References

J. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer, New York, 1997.

J. S. Cundiff, N. Dias, and H. D. Sherali. A linear programming approach for designing a herbaceous biomass delivery system. *Bioresource technology*, 59(1):47–55, 1997.

A. De Meyer, C. Dirk, J. Rasinmäki, and J. Van Orshoven. Methods to optimise the design and management of biomass-for-bioenergy supply chains: A review. *Renewable and Sustainable Energy Reviews*, 31:657–670, 2014. ISSN 1364-0321. doi: 10.1016/j.rser.2013.12.036.

S. van Dyken, B. H. Bakken, and H. I. Skjelbred. Linear mixed-integer models for biomass supply chains with transport, storage and processing. *Energy*, 35(3):1338–1350, 2010. doi: 10.1016/j.energy.2009.11.017.

ecoprog GmbH. Biomass to energy 2013/2014 – The world market for biomass power plants. Technical report, 2013. URL `http://www.ecoprog.com/en/publications/energy-industry/biomass-to-energy.htm`.

H. Gunnarsson, M. Rönnqvist, and J. T. Lundgren. Supply chain modelling of forest fuel. *European Journal of Operational Research*, 158(1):103–123, 2004. ISSN 0377-2217. doi: http://dx.doi.org/10.1016/S0377-2217(03)00354-0.

C. Kanzian, F. Holzleitner, K. Stampfer, and S. Ashton. Regional energy wood logistics – optimizing local fuel supply. *Silva Fennica*, 43(1):113–128, 2009. URL `http://www.metla.fi/silvafennica/full/sf43/sf431113.pdf`.

F. Mafakheri and F. Nasiri. Modeling of biomass-to-energy supply chain operations: Applications, challenges and research directions. *Energy Policy*, 2013.

D. Rizzo, L. Martin, and J. Wohlfahrt. Miscanthus spatial location as seen by farmers: a machine learning approach to model real criteria. *Biomass & Bioenergy*, in press 2014.

J. A. Scott, W. Ho, and P. K. Dey. A review of multi-criteria decision-making methods for bioenergy systems. *Energy*, 42(1):146–156, 2012. ISSN 0360-5442. doi: http://dx.doi.org/10.1016/j.energy.2012.03.074. URL `http://www.sciencedirect.com/science/article/pii/S0360544212002770`. 8th World Energy System Conference, {WESC} 2010.

N. Shabani, S. Akhtari, and T. Sowlati. Value chain optimization of forest biomass for bioenergy production: A review. *Renewable and Sustainable Energy Reviews*, 23: 299–311, 2013. doi: 10.1016/j.rser.2013.03.005.

B. Sharma, R. Ingalls, C. Jones, and A. Khanchi. Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, 24:608–627, 2013.

B. K. Thapalia, S. W. Wallace, and M. Kaut. Using inventory to handle risks in the supply of oil to Nepal. *International Journal of Business Performance and Supply Chain Modelling*, 1(1):41–60, 2009. ISSN 1758-941X (online), 1758-9401 (print). doi: 10.1504/IJBPSCM.2009.026265.

X. van Tilburg, R. Egging, and H. Londo. Biotrans functional and technical description. ECN Policy Studies ECN-RX–06-013, Energy research Centre of the Netherlands, 2006. URL `https://www.ecn.nl/publications/ECN-RX--06-013`.

U.S. Energy Information Agency. International energy outlook 2013. Technical Report DOE/EIA-0484(2013), 2013. URL `http://www.eia.gov/forecasts/ieo/`.

G. Walther, A. Schatka, and T. S. Spengler. Design of regional production networks for second generation synthetic bio-fuel–a case study in northern germany. *European Journal of Operational Research*, 218(1):280–292, 2012.

U. J. Wolfsmayr and P. Rauch. The primary forest fuel supply chain: A literature review. *Biomass and Bioenergy*, 60:203–221, 2014.

D. Yue, F. You, and S. W. Snyder. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, In press, 2014. doi: 10.1016/j.compchemeng.2013.11.016.